

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/763,249  
Source: IFWO  
Date Processed by STIC: 10/18/04

# ***ENTERED***



IFWO

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/763,249

TIME: 09:52:17

Input Set : A:\247264US0CONT.txt

Output Set: N:\CRF4\09282004\J763249.raw

3 <110> APPLICANT: TONOUCHI, NAOTO  
 4 SUZUKI, SONOKO  
 5 YOKOZEKI, KENZO  
 6 NOZAKI, HIROYUKI  
 7 SUGIYAMA, MASAKAZU  
 9 <120> TITLE OF INVENTION: PEPTIDE-FORMING ENZYME GENE, PEPTIDE-FORMING ENZYME, AND

## PEPTIDE

10 PRODUCING METHOD  
 12 <130> FILE REFERENCE: 247264US0CONT  
 14 <140> CURRENT APPLICATION NUMBER: 10/763,249  
 15 <141> CURRENT FILING DATE: 2004-01-26  
 17 <150> PRIOR APPLICATION NUMBER: PCT/JP02/07635  
 18 <151> PRIOR FILING DATE: 2002-07-26  
 20 <150> PRIOR APPLICATION NUMBER: JP 2001-226568  
 21 <151> PRIOR FILING DATE: 2001-07-26  
 23 <150> PRIOR APPLICATION NUMBER: JP 2001-310547  
 24 <151> PRIOR FILING DATE: 2001-10-05  
 26 <160> NUMBER OF SEQ ID NOS: 17  
 28 <170> SOFTWARE: PatentIn version 3.3  
 30 <210> SEQ ID NO: 1  
 31 <211> LENGTH: 30  
 32 <212> TYPE: PRT  
 33 <213> ORGANISM: Corynebacterium glutamicum  
 36 <220> FEATURE:  
 37 <221> NAME/KEY: misc\_feature  
 38 <222> LOCATION: (26)..(26)  
 39 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 41 <220> FEATURE:  
 42 <221> NAME/KEY: misc\_feature  
 43 <222> LOCATION: (28)..(28)  
 44 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 46 <220> FEATURE:  
 47 <221> NAME/KEY: misc\_feature  
 48 <222> LOCATION: (30)..(30)  
 49 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid  
 51 <400> SEQUENCE: 1  
 53 Thr Lys Thr Leu Gly Ser Leu Gln Leu Glu Glu Ile Thr Leu Thr Leu  
 54 1 5 10 15  
 W--> 57 Pro Leu Thr Glu Asp Val Ala Asp Glu Xaa Arg Xaa Glu Xaa  
 58 20 25 30

61 <210> SEQ ID NO: 2  
 62 <211> LENGTH: 26  
 63 <212> TYPE: DNA  
 64 <213> ORGANISM: Artificial Sequence

P.6

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/763,249

TIME: 09:52:17

Input Set : A:\247264US0CONT.txt

Output Set: N:\CRF4\09282004\J763249.raw

```

66 <220> FEATURE:
67 <223> OTHER INFORMATION: Synthetic DNA
70 <220> FEATURE:
71 <221> NAME/KEY: misc_feature
72 <222> LOCATION: (6)..(6)
73 <223> OTHER INFORMATION: n is a, c, g, or t
75 <400> SEQUENCE: 2
W--> 76 gghwsnytbcb arytbgarga ratyac 26
79 <210> SEQ ID NO: 3
80 <211> LENGTH: 27
81 <212> TYPE: DNA
82 <213> ORGANISM: Artificial Sequence
84 <220> FEATURE:
85 <223> OTHER INFORMATION: Synthetic DNA
87 <400> SEQUENCE: 3
88 carytbgarg aratyacbyt bacbytb 27
91 <210> SEQ ID NO: 4
92 <211> LENGTH: 1307
93 <212> TYPE: DNA
94 <213> ORGANISM: Corynebacterium glutamicum
97 <220> FEATURE:
98 <221> NAME/KEY: CDS
99 <222> LOCATION: (57)..(1295)
101 <400> SEQUENCE: 4
102 ggcgagctcg ggcagtggtg ggggtggtgt ccaccctgc gcgtaacctg ggaagc atg 59
103 Met
104 1
106 act aaa aca ctt ggt tcc ctt caa ctt gaa gaa att acc ttg acg ctc 107
107 Thr Lys Thr Leu Gly Ser Leu Gln Leu Glu Glu Ile Thr Leu Thr Leu
108 5 10 15
110 cct ctg act gaa gat gtg gcc gat gaa cgc acc att gat gtg ttc gca 155
111 Pro Leu Thr Glu Asp Val Ala Asp Glu Arg Thr Ile Asp Val Phe Ala
112 20 25 30
114 cgc att gcc aca cgc gtc ggt ggg gaa gac ctt cca tat tta gta ttc 203
115 Arg Ile Ala Thr Arg Val Gly Gly Glu Asp Leu Pro Tyr Leu Val Phe
116 35 40 45
118 ctg cag ggt ggg cct ggc aat gaa gct cca cgt cca agc ctt aat ccc 251
119 Leu Gln Gly Gly Pro Gly Asn Glu Ala Pro Arg Pro Ser Leu Asn Pro
120 50 55 60 65
122 ctc aac ccc aat tgg ttg ggc gtg gcc ttg gag gaa tac cgc gtg gtc 299
123 Leu Asn Pro Asn Trp Leu Gly Val Ala Leu Glu Glu Tyr Arg Val Val
124 70 75 80
126 atg ttg gat caa cgt ggc acc ggc cgt tcc acc cca gtg ggt aat gat 347
127 Met Leu Asp Gln Arg Gly Thr Gly Arg Ser Thr Pro Val Gly Asn Asp
128 85 90 95
130 att ttg gaa aaa ccc aca gca gaa gta gtg gag tac tta tcc cac ctg 395
131 Ile Leu Glu Lys Pro Thr Ala Glu Val Val Glu Tyr Leu Ser His Leu
132 100 105 110
134 cgc gca gat ggc att gtg cga gat gct gaa gcc ctg cgt aag cat ttg 443

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/763,249

TIME: 09:52:17

Input Set : A:\247264US0CONT.txt

Output Set: N:\CRF4\09282004\J763249.raw

135	Arg	Ala	Asp	Gly	Ile	Val	Arg	Asp	Ala	Glu	Ala	Leu	Arg	Lys	His	Leu	
136		115					120					125					
138	ggt	gtg	aat	cag	tgg	aac	ctt	tta	ggc	cag	tcc	ttc	gga	ggt	ttc	acc	491
139	Gly	Val	Asn	Gln	Trp	Asn	Leu	Leu	Gly	Gln	Ser	Phe	Gly	Gly	Phe	Thr	
140	130					135					140					145	
142	acc	ctg	cat	tac	ttg	tcc	cgg	cac	gcc	gat	tcc	ttg	gac	aac	gtg	ttt	539
143	Thr	Leu	His	Tyr	Leu	Ser	Arg	His	Ala	Asp	Ser	Leu	Asp	Asn	Val	Phe	
144					150					155					160		
146	att	acc	ggc	ggt	ctc	agc	gct	att	gat	cgc	cca	gca	gaa	gac	gtg	tat	587
147	Ile	Thr	Gly	Gly	Leu	Ser	Ala	Ile	Asp	Arg	Pro	Ala	Glu	Asp	Val	Tyr	
148			165						170						175		
150	gcc	aac	tgt	tac	aac	cgc	atg	cgc	cga	aac	tct	gag	gaa	ttc	tac	cgt	635
151	Ala	Asn	Cys	Tyr	Asn	Arg	Met	Arg	Arg	Asn	Ser	Glu	Glu	Phe	Tyr	Arg	
152		180					185					190					
154	cgc	ttc	ccg	caa	tta	cgg	gaa	act	ttc	cga	ggg	ttg	gtt	aat	cgt	gct	683
155	Arg	Phe	Pro	Gln	Leu	Arg	Glu	Thr	Phe	Arg	Gly	Leu	Val	Asn	Arg	Ala	
156		195				200					205						
158	cgc	gcc	ggg	gag	att	gtg	ctt	ccc	acc	ggc	gaa	gtt	gtg	tca	gaa	acc	731
159	Arg	Ala	Gly	Glu	Ile	Val	Leu	Pro	Thr	Gly	Glu	Val	Val	Ser	Glu	Thr	
160	210				215						220				225		
162	agg	ctg	cga	tcc	ctt	ggt	cac	ttg	ttg	ggt	agc	aat	gac	ggc	tgg	ttt	779
163	Arg	Leu	Arg	Ser	Leu	Gly	His	Leu	Leu	Gly	Ser	Asn	Asp	Gly	Trp	Phe	
164			230							235					240		
166	gat	ctg	tac	aac	ctg	ctg	gaa	tta	gat	ccc	acc	tcc	aac	gct	ttt	gtc	827
167	Asp	Leu	Tyr	Asn	Leu	Leu	Glu	Leu	Asp	Pro	Thr	Ser	Asn	Ala	Phe	Val	
168			245						250					255			
170	cat	gac	ctg	gca	gga	ctt	ttg	cct	ttc	ggc	aac	cgc	aac	cca	att	tat	875
171	His	Asp	Leu	Ala	Gly	Leu	Leu	Pro	Phe	Gly	Asn	Arg	Asn	Pro	Ile	Tyr	
172		260					265					270					
174	tac	gtg	ctc	cat	gag	tcc	tct	tac	gcc	gac	ggt	gtg	gtg	aca	aat	tgg	923
175	Tyr	Val	Leu	His	Glu	Ser	Ser	Tyr	Ala	Asp	Gly	Val	Val	Thr	Asn	Trp	
176		275				280					285						
178	gca	gca	gag	cgt	gtg	ctt	cca	gag	gat	ttc	cgc	gag	gat	cca	aca	ctg	971
179	Ala	Ala	Glu	Arg	Val	Leu	Pro	Glu	Asp	Phe	Arg	Glu	Asp	Pro	Thr	Leu	
180	290				295					300					305		
182	ctc	acc	ggt	gag	cac	gtg	ttc	cag	gag	tgg	aca	gac	acc	gtg	ccg	tgc	1019
183	Leu	Thr	Gly	Glu	His	Val	Phe	Gln	Glu	Trp	Thr	Asp	Thr	Val	Pro	Ser	
184			310							315					320		
186	ctc	aag	ccg	tgg	aag	gac	gtt	gcc	ctg	gca	ttg	gct	cag	cag	gaa	tgg	1067
187	Leu	Lys	Pro	Trp	Lys	Asp	Val	Ala	Leu	Ala	Leu	Ala	Gln	Gln	Glu	Trp	
188			325						330					335			
190	ccc	aag	ctt	tat	gat	gcg	aag	gca	ttg	gaa	aac	tca	cag	gcc	aag	ggc	1115
191	Pro	Lys	Leu	Tyr	Asp	Ala	Lys	Ala	Leu	Glu	Asn	Ser	Gln	Ala	Lys	Gly	
192		340					345					350					
194	gct	gca	gca	gtg	tat	ghc	aat	gac	gtt	ttc	gtc	cca	gtg	gat	tac	tct	1163
W--> 195	Ala	Ala	Ala	Val	Tyr	Xaa	Asn	Asp	Val	Phe	Val	Pro	Val	Asp	Tyr	Ser	
196		355					360					365					
198	ctg	gaa	acc	gca	caa	cac	ctg	ccc	ggt	gtg	cag	ctg	ttt	atc	acc	agc	1211
199	Leu	Glu	Thr	Ala	Gln	His	Leu	Pro	Gly	Val	Gln	Leu	Phe	Ile	Thr	Ser	

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/763,249

TIME: 09:52:17

Input Set : A:\247264US0CONT.txt

Output Set: N:\CRF4\09282004\J763249.raw

```

200 370          375          380          385
202 cag cat gaa cac aat gga ctt cgt gcc agc tca ggc gca gta ctg rag      1259
203 Gln His Glu His Asn Gly Leu Arg Ala Ser Ser Gly Ala Val Leu Xaa
204          390          395          400
206 cac ctt ttc gat ctg gcc cac ggc cga gag gta cgc tgagggcccc cg      1307
207 His Leu Phe Asp Leu Ala His Gly Arg Glu Val Arg
208          405          410
211 <210> SEQ ID NO: 5
212 <211> LENGTH: 413
213 <212> TYPE: PRT
214 <213> ORGANISM: Corynebacterium glutamicum
216 <220> FEATURE:
217 <221> NAME/KEY: misc_feature
218 <222> LOCATION: (359)..(359)
219 <223> OTHER INFORMATION: The 'Xaa' at location 359 stands for Asp, Ala, or Val.
221 <220> FEATURE:
222 <221> NAME/KEY: misc_feature
223 <222> LOCATION: (401)..(401)
224 <223> OTHER INFORMATION: The 'Xaa' at location 401 stands for Glu, or Lys.
226 <400> SEQUENCE: 5
228 Met Thr Lys Thr Leu Gly Ser Leu Gln Leu Glu Glu Ile Thr Leu Thr
229 1          5          10          15
232 Leu Pro Leu Thr Glu Asp Val Ala Asp Glu Arg Thr Ile Asp Val Phe
233          20          25          30
236 Ala Arg Ile Ala Thr Arg Val Gly Gly Glu Asp Leu Pro Tyr Leu Val
237          35          40          45
240 Phe Leu Gln Gly Gly Pro Gly Asn Glu Ala Pro Arg Pro Ser Leu Asn
241          50          55          60
244 Pro Leu Asn Pro Asn Trp Leu Gly Val Ala Leu Glu Glu Tyr Arg Val
245 65          70          75          80
248 Val Met Leu Asp Gln Arg Gly Thr Gly Arg Ser Thr Pro Val Gly Asn
249          85          90          95
252 Asp Ile Leu Glu Lys Pro Thr Ala Glu Val Val Glu Tyr Leu Ser His
253          100         105         110
256 Leu Arg Ala Asp Gly Ile Val Arg Asp Ala Glu Ala Leu Arg Lys His
257          115         120         125
260 Leu Gly Val Asn Gln Trp Asn Leu Leu Gly Gln Ser Phe Gly Gly Phe
261          130         135         140
264 Thr Thr Leu His Tyr Leu Ser Arg His Ala Asp Ser Leu Asp Asn Val
265 145         150         155         160
268 Phe Ile Thr Gly Gly Leu Ser Ala Ile Asp Arg Pro Ala Glu Asp Val
269          165         170         175
272 Tyr Ala Asn Cys Tyr Asn Arg Met Arg Arg Asn Ser Glu Glu Phe Tyr
273          180         185         190
276 Arg Arg Phe Pro Gln Leu Arg Glu Thr Phe Arg Gly Leu Val Asn Arg
277          195         200         205
280 Ala Arg Ala Gly Glu Ile Val Leu Pro Thr Gly Glu Val Val Ser Glu
281          210         215         220
284 Thr Arg Leu Arg Ser Leu Gly His Leu Leu Gly Ser Asn Asp Gly Trp

```

## RAW SEQUENCE LISTING

DATE: 10/18/2004

PATENT APPLICATION: US/10/763,249

TIME: 09:52:17

Input Set : A:\247264US0CONT.txt

Output Set: N:\CRF4\09282004\J763249.raw

```

285 225          230          235          240
288 Phe Asp Leu Tyr Asn Leu Leu Glu Leu Asp Pro Thr Ser Asn Ala Phe
289          245          250          255
292 Val His Asp Leu Ala Gly Leu Leu Pro Phe Gly Asn Arg Asn Pro Ile
293          260          265          270
296 Tyr Tyr Val Leu His Glu Ser Ser Tyr Ala Asp Gly Val Val Thr Asn
297          275          280          285
300 Trp Ala Ala Glu Arg Val Leu Pro Glu Asp Phe Arg Glu Asp Pro Thr
301          290          295          300
304 Leu Leu Thr Gly Glu His Val Phe Gln Glu Trp Thr Asp Thr Val Pro
305 305          310          315          320
308 Ser Leu Lys Pro Trp Lys Asp Val Ala Leu Ala Leu Ala Gln Gln Glu
309          325          330          335
312 Trp Pro Lys Leu Tyr Asp Ala Lys Ala Leu Glu Asn Ser Gln Ala Lys
313          340          345          350
W--> 316 Gly Ala Ala Ala Val Tyr Xaa Asn Asp Val Phe Val Pro Val Asp Tyr
317          355          360          365
320 Ser Leu Glu Thr Ala Gln His Leu Pro Gly Val Gln Leu Phe Ile Thr
321          370          375          380
324 Ser Gln His Glu His Asn Gly Leu Arg Ala Ser Ser Gly Ala Val Leu
325 385          390          395          400
328 Xaa His Leu Phe Asp Leu Ala His Gly Arg Glu Val Arg
329          405          410
332 <210> SEQ ID NO: 6
333 <211> LENGTH: 30
334 <212> TYPE: DNA
335 <213> ORGANISM: Artificial Sequence
337 <220> FEATURE:
338 <223> OTHER INFORMATION: Synthetic DNA
340 <400> SEQUENCE: 6
341 ggcgagctcg ggcagtgggtg ggggtgggtgt 30
344 <210> SEQ ID NO: 7
345 <211> LENGTH: 30
346 <212> TYPE: DNA
347 <213> ORGANISM: Artificial Sequence
349 <220> FEATURE:
350 <223> OTHER INFORMATION: Synthetic DNA
352 <400> SEQUENCE: 7
353 cgggggccct cagcgtacct ctgggccgtg 30
356 <210> SEQ ID NO: 8
357 <211> LENGTH: 30
358 <212> TYPE: DNA
359 <213> ORGANISM: Artificial Sequence
361 <220> FEATURE:
362 <223> OTHER INFORMATION: Synthetic DNA
364 <400> SEQUENCE: 8
365 ggcgagctca tgactaaaac acttggttcc 30
368 <210> SEQ ID NO: 9
369 <211> LENGTH: 25

```

RAW SEQUENCE LISTING ERROR SUMMARY  
PATENT APPLICATION: US/10/763,249

DATE: 10/18/2004  
TIME: 09:52:18

Input Set : A:\247264US0CONT.txt  
Output Set: N:\CRF4\09282004\J763249.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 26,28,30  
Seq#:2; N Pos. 6  
Seq#:4; Xaa Pos. 359,401  
Seq#:5; Xaa Pos. 359,401  
Seq#:16; N Pos. 1297,1298

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/763,249

DATE: 10/18/2004

TIME: 09:52:18

Input Set : A:\247264US0CONT.txt

Output Set: N:\CRF4\09282004\J763249.raw

L:57 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:16  
L:76 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0  
L:195 M:258 W: Mandatory Feature missing, <223> Tag not found for SEQ ID#:4  
L:195 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:1163  
M:341 Repeated in SeqNo=4  
L:316 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:352  
M:341 Repeated in SeqNo=5  
L:763 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16 after pos.:1261